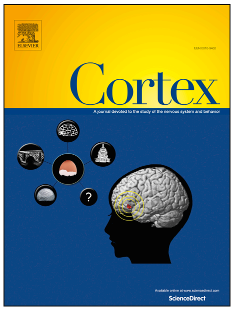
On the importance of cognitive profiling

Massa, M. Sofia; Wang, Naxian; Bickerton, Wai-Ling; Demeyere, Nele; Riddoch, M. Jane; Humphreys, Glyn W.

DOI:
10.1016/j.cortex.2015.06.006

Accepted Manuscript



On the importance of cognitive profiling: A graphical modelling analysis of domain-specific and domain-general deficits after stroke

M.S. Massa, N. Wang, W.-L. Bickerton, N. Demeyere, M.J. Riddoch, G.W. Humphreys

PII: S0010-9452(15)00208-7

DOI: 10.1016/j.cortex.2015.06.006

Reference: CORTEX 1502

To appear in: *Cortex*

Received Date: 3 February 2014

Revised Date: 8 November 2014

Accepted Date: 10 June 2015

On the importance of cognitive profiling: A graphical modelling analysis of domain-specific and domain-general deficits after stroke

Massa, M.S.[1], Wang, N.[1], Bickerton, W-L.[2], Demeyere, N.[3], Riddoch, M.J.[3] & Humphreys, G.W.[3]

[1]Department of Statistics

Oxford University

Oxford OX1 3UD, UK

[2]School of Psychology

University of Birmingham

Birmingham B15 2TT, UK

[3]Department of Experimental Psychology

Oxford University

Oxford OX1 3UD, UK

Running head: Cognitive screening of domain-specific and general deficits

Abstract

Cognitive problems following stroke are typically analysed using either short but relatively uninformative general tests or through detailed but time consuming tests of domain specific deficits (e.g., in language, memory, praxis). Here we present an analysis of neuropsychological deficits detected using a screen designed to fall between other screens by being ‘broad’ (testing multiple cognitive abilities) but ‘shallow’ (sampling the abilities briefly, to be time efficient) – the BCoS. Assessment using the BCoS enables the relations between ‘domain specific’ and ‘domain general’ cognitive deficits to be evaluated as the test generates an overall cognitive profile for individual patients. We analysed data from 287 patients tested at a sub-acute stage of stroke (<3 months). Graphical modelling techniques were used to investigate the associative structure and conditional independence between deficits within and across the domains sampled by BCoS (attention and executive functions, language, memory, praxis and number processing). The patterns of deficit within each domain conformed to existing cognitive models. However, these within-domain patterns underwent substantial change when the whole dataset was modelled, indicating that domain-specific deficits can only be understood in relation to linked changes in domain-general processes. The data point to the importance of using over-arching cognitive screens, measuring domain-general as well as domain-specific processes, in order to account for neuropsychological deficits after stroke. The paper also highlights the utility of using graphical modelling to understand the relations between cognitive components in complex datasets.

INTRODUCTION

Cognitive problems are highly prevalent after stroke, occurring in up to 70% of the sub-acute stroke population (Humphreys et al., 2012; Nys et al., 2005). Furthermore, these problems are strongly predictive of poor recovery of function, even after the presence of motor problems have been taken into account (e.g., see Bickerton et al., 2011, 2012; de Haan, Nys, & Van Zandvoort, 2006; Narasimhalu et al., 2009; Nys et al., 2005; van Zandvoort, Kessels, Nys, de Haan, & Kappelle, 2005). Given their importance for predicting outcome, and given the costs of poor recovery, it is critical that cognitive deficits are diagnosed early so that diagnosis can inform rehabilitation.

Current attempts to screen for cognitive deficits after stroke typically take one of two forms. There exist several short screens which can be administered relatively easily but can be relatively unspecific in diagnosing domain-specific deficits. Examples here include the mini-mental state examination (MMSE), the MOCA and the ACE-R. One limitation of these tests is that they were designed to aid the diagnosis of dementia. As cognitive problems following stroke differ from those in dementia, the tests are insensitive to particular deficits after stroke (Demeyere et al., in press). For example, apraxia and neglect are both relatively common consequences of stroke (affecting above 30% of the left and right hemisphere populations; Humphreys et al., 2012) and both predict longer-term outcome (Bickerton et al., 2011, 2012), but neither are systematically detected by these screens. Contrasting with these short screens are neuropsychological tests such as PALPA (Kay, Lesser & Coltheart, 1993), the Behavioural Inattention Test (Wilson, Cockburn & Halligan, 1987), the Doors and Peoples test (Baddeley, Emslie & Nimmo-Smith, 1994) and so forth. These tests analyse deficits in different cognitive domains (respectively language, attention and memory in our examples) but require administration times that are too long for many clinical settings. These assessments also encourage a focus on the specific cognitive domain the test examines.

In order to overcome these problems, we (Humphreys et al., 2012) developed the BCoS (Birmingham Cognitive Screen). The BCoS is designed to sample domain-specific deficits in 5 areas of cognition (attention and executive function, language, memory, number abilities and praxis) but in a relatively time efficient manner – taking about 1 hour to administer. Test scores can be used to diagnose the 'cognitive profile' of a patient across several cognitive domains, measured relative to age and education match control data, which can be used in clinical case management (Humphreys et al., 2012). For clinical case management, the cognitive profile is reported (see Appendix 1).

One other attribute of BCoS is that it is designed to (i) measure some of the deficits frequently found after stroke (including spatial neglect, poor number processing, apraxia) and (ii) the tests are designed so as not to be confounded by the presence of some common post-stroke problems (e.g., when the tests do not aim to measure language or spatial attention are affected by the presence of aphasia or unilateral neglect). These design attributes distinguish BCoS from some of the other general cognitive screens in the field (e.g., the MMSE, MOCA), which do not measure some of the critical deficits and which are largely language dependent and which can be affected by spatial neglect.

Domain-specific and domain-general factors.

The domain-specific analysis of cognitive problems after brain injury has historically characterised neuropsychological assessment and matches the treatment of neuropsychological problems in clinical textbooks (Andrewes, 2001; Heilman & Valenstein, 2011; Rapp, 2001). However, there are grounds for doubting that domain-specific analyses are sufficient for giving an appropriate characterisation of patients. For example, within the domain of spatial attention, there is considerable evidence indicating that clinical deficits such as unilateral neglect are greatly exacerbated if patients have ancillary deficits in sustained attention and working memory (Malhotra et al., 2005). Similarly, language problems in patients are increased when executive and working memory deficits are present, which can disrupt a patient's ability to control lexical selection and to maintain phonological codes during sentence processing (Brownsett et al., 2014; Fillingham et al., 2005; Francis, Clark & Humphreys, 2003). Corbetta et al. (2015) have also recently argued that the variance across stroke patients can be substantially captured by three domain-general factors covering: (i) language, verbal and spatial memory, (ii) left-side motor weakness, right visual field bias and attention shifting, and (iii) right side motor weakness and left visual field bias. They note that deficits across multiple domains are associated with damage to 'cross road' regions of white matter, where multiple white matter tracts are present. Co-occurring impairments in different cognitive processes can have an impact on and modulate the expression of what are typically treated as domain-specific deficits. Screens such as BCoS can offer a different approach to the discrete, domain-specific analysis of cognitive deficits, since the screens emphasise the 'cognitive profile' for a patient and include measures of co-occurring problems in different domains alongside any domain-specific cognitive impairments (e.g., measures of sustained attention and working memory are taken along with measures of language; see Appendix 1). This then makes screens such as BCoS sensitive to the interaction between what we will term 'domain-general deficits' (e.g., working memory and sustained attention, which are required to support processing in several different domains such as language, spatial attention etc.) and what are putatively domain-specific impairments (language, memory, spatial orienting of attention). Here we evaluate whether the cognitive profiling approach, promoted by BCoS, can provide new insights into the nature of domain-specific deficits when domain-general processes are taken into account. We report data from a large-scale screening programme of cognitive problems after stroke, conducted using the BCoS.

To bring out our argument about the contribution of domain-general as well as domain-specific processes we introduce a relatively new way of analysing data from large datasets involving multiple different cognitive tests: graphical models analysis. Traditional approaches to analysing data from neuropsychological test batteries have conducted factor or cluster analyses (e.g., see Corbetta et al., 2015; Miyake et al., 2000; Verdon et al., 2010). Each of these approaches has its virtues – for example, factor analysis is useful for bringing out hidden factors which may contribute to several domains and for assessing which tests combine together to generate particular results. Graph analyses go beyond these other approaches, however, by testing the conditional independence of different assessments directly, without making assumptions about underlying hidden factors. Moreover, graphical models tell us more than cluster analysis because it can provide information about the strength of the links between different variables, whereas clustering only creates groups of similar variables. Graphical models also go beyond multiple regression approaches which assess the linear dependency of a measure on some explanatory variables; in graphical model analysis we capture the interactions between all the variables. Here the analysis evaluated the relations between the different sub-tests of the BCoS when patient performance was considered both at (i) a domain-specific level (considering language separately from spatial attention and so forth) and when (ii) domain-general measures could contribute, when all the tests were considered together. We assessed whether the data patterns that emerged between the tests when they were analysed within a domain, as is standardly done, were substantially changed when a domain-general analysis was undertaken, taking all of the tests into account. Performance on the BCoS was also evaluated in relation to measures of motor performance (the Barthel index; Mahoney & Barthel, 1965) and affect (the Hospital Anxiety and Depression scale; Zigmond & Snaith, 1983), to ensure that changes introduced by a domain-general analysis were not confounded by co-varying problems outside of cognition (in motor function or affect, captured by the Barthel and HADS scores respectively).

Graphical model analysis

Graphical modelling is a form of multivariate analysis that originated from the fields of physics (Gibbs, 1902) and genetics (Wright, 1921) (See Højsgaard et al., 2012, for an overview on the methods and their implementation within the statistical software R, available at http://cran.r-project.org/). Graphical models combine the notion of a statistical model with a mathematical object, a graph. In particular, given a study of interest, every random variable is represented via a vertex (node) in a graph. The nodes can be connected by different types of 'edges' (which may be undirected or directed), reflecting the statistical relations between the variables – in this case the sub-tests of the BCoS. In this paper we will focus on undirected graphical models, i.e., graphical models where graphs have only undirected edges (not assuming directed relations between any two tests). Undirected edges represent associations between random variables and a missing edge reflects the fact that random variables are conditionally independent.

The notion of conditional independence is important for understanding graphical modelling. Given three random variables X, Y and Z, X and Y are conditionally independent given Z, if for each value z of Z, X and Y are independent in the conditional distribution given $Z = z$. Essentially if the value taken by Z is known, information about Y is irrelevant for knowledge of X while information about X is irrelevant for knowledge of Y. For a continuous distribution, this is equivalent to saying that the joint distribution of the three random variables can be factorised as

$$f_{X,Y,Z}(x, y, z) = \frac{f_{X,Z}(x,z) f_{Y,Z}(y,z)}{f_Z(z)}.$$

As an illustration, consider the example of a study of health and social characteristics of 70-year-olds taken at two intervals (e.g., in 1967 and again in 1984; see Edwards 2000). Edwards refers to the distribution of body mass index (BMI) between males and females, and between the two years of sampling. If males and females have differing distributions of BMI, but there has been no change in these distributions across time, then the variable BMI is conditionally independent of the variable Year, given the variable Gender: essentially, if we know Gender, information about Year is irrelevant for knowledge of BMI. On the other hand, if males and females have the same distribution of BMI, but this changes from 1967 to 1984 then BMI is conditionally independent of Gender given Year. In this case, if we know Year then information about Gender is irrelevant for knowledge of BMI. The conditional dependence between any two factors can also change as other factors are added into the model. For example, the conditional dependence between BMI and Year may reduce if, across the years, individuals differed in height and height is included in the full model analysis.

The key tool in graphical modelling is the dependence graph. A graph, denoted by $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ consists of a finite set V of vertices and a finite set E of edges. In a dependence graph the vertices represent random variables of a multivariate distribution, and two vertices either have one edge or no edge between them and the missing edges represent conditional independences between the random variables in the following way. If two vertices A and B are separated in the graph by a vertex (or a set of vertices) S, then the corresponding random variables A and B are conditionally independent given S. Hence conditional independence relations are directly read off the graph. This is the so called global Markov property and it establishes the correspondence between nodes in a graph and conditional independence relations between the variables of the multivariate distribution under consideration (for a detailed exposition of Markov properties and their equivalence, see Edwards, 2000; Whittaker, 1990 or Lauritzen, 1996).

Over the past 20 years, the development of graphical modelling has enabled researchers to explore the complex structure of high-dimensional data using both visually and computationally powerful tools. This analysis of the structural relations between tests enables graph modelling to extend traditional approaches in neuropsychology, where the relations between the sub-tests are not precisely specified. Our objective was to apply graphical modelling techniques to the BCoS data set, to assess the relations between domain-specific and domain-general analyses and to test whether the relations between the different cognitive measures hold even when variance reflecting clinical deficits in motor abilities and affect is extracted.

We report the data in two parts. First, we analysed the structure of the results within each of the putative domains of the BCoS. This provides the kind of domain-specific analysis that is derived when researchers focus on one aspect of cognition such as language or memory. Second, we analysed the structure of the results when all of the test domains were reviewed together, to examine the extent to which the domain-specific organisation remained when performance on the other tests and domains was taken into account. Using the graph modelling approach we ask whether the structural relations between tests changes when cross-domain data are included?

Methods

The dataset

The data set contained the cognitive profile of 287 stroke patients with complete BCoS scores[1]. There were 41 variables in the data set with 4 personal information variables, 9 clinical information variables and 28 test variables (the cognitive test scores). The personal information variables included age, gender, education level, and handedness.

The clinical information recorded was divided into physical and behavioural

[1] Researchers interested in accessing the dataset should contact Glyn Humphreys: glyn.humphreys@psy.ox.ac.uk

variables. The physical variables included the patient's stroke history (previous stroke, TIA, head injury, dementia etc.), the type of stroke (haemorrhagic, ischaemic), the visible presence of a lesion on CT scan, the side of the lesion (left, right), the lesion location (cortical, subcortical) and whether more general vascular changes were noted. Behavioural variables included information about the general physical and psychological condition of the patients with scores generated using the Barthel ADL Index and the Hospital Anxiety and Depression Scale (HADS).

The six physical variables reflected factors that could exert a direct impact on patients' test performance. Jørgensen et al. (1995) showed that the type of stroke could determine the severity of the patients' conditions. In addition, cognitive abilities are likely to deteriorate more if patients suffer from repeated strokes (Bickerton et al., in press; Jørgensen et al., 1997). Variance due to these clinical factors was taken into account within the graph analyses. In addition, we took two behavioural clinical measures reflecting the patients' general physical and affective condition. The Barthel ADL Index (Mahoney & Barthel, 1965) provides a measure of performance in activities of daily living (ADL)(score 0-20, where a high score = more able). The HADS (Zigmond & Snaith, 1983) provides measures of anxiety and depression, with higher scores indicating increased symptoms. We included these physical and affective measures within the full graph analysis reported below, so we could depict their relation with the cognitive measures provided by the BCoS.

The central information of the BCoS dataset is the performance of the patients in different cognitive sub-tests. There are 23 sub-tests in total (Humphreys et al., 2012), with between 3 and 5 cognitive tests in each putative domain. The majority of the test variables measure the absolute level of abilities with a higher score standing for better performance. In some cases, test measures reflect relative differences between conditions – examples being the relative performance on left and right side stimuli in the measures of unilateral neglect (the Apple

Cancellation Asymmetry score) and extinction (in the visual and tactile extinction tasks). Higher scores in these tests stand for a stronger asymmetry rather than better performance. Full details of the tests making up the BCoS are reported in Humphreys et al. (2012) and short descriptions are given in Appendices 1 and 2.

Results

Patients' personal traits

Male patients slightly outnumbered female patients (176 vs. 111). About three quarters (74%) of the patients had secondary school education as their highest education attainment and 16% graduated from college. 3% had a non-university diploma, and 7% went to universities. A small proportion of patients (2%) only had primary school education. 89.2% of the patients self-reported as right-handed, with 8.7% reporting as being pre-morbidly left-handed and 2.1% as ambidextrous.

80.1% of patients had ischemic stroke and a minority intra-cerebral haemorrhage (18.1%), confirmed on CT scan. 58% had a unilateral right side lesion and 42% a unilateral left-side lesion[2]. Only cortical damage was noted in 34.8% of the population, while 35.5% had sub-cortical lesions followed by the subcortical region and 11.5% of the patients suffered from grey and white matter lesions.

Structural learning on domain-specific tests

When working with graphical models, the process of selecting a model that best fits the data is called structural learning because the aim is to infer the structure (the dependence graph) that best describes the conditional independences and associations between the random

[2] Patients with bilateral lesions or without a lesion confirmed on CT scan were omitted from the data analysis.

variables. Based on the dependence graph obtained we can then obtain estimates for the model parameters. In Section 1 we used model selection strategies for continuous variables based on sub-sets of the BCoS dataset. We looked separately at models for the clinical behavioural variables and the cognitive test variables in Section 2. Given the moderate size of each subset, we investigated graphical models stepwise procedures using the Bayesian Information Criterion (BIC) (Schwarz, 1978) (see Højsgaard et al., 2012, for details). The idea was to start from an initial model (e.g., the complete independence model, with no edges between nodes) and at each step add or delete the edge that gives the largest decrease in the significance testing via BIC. If there is no change in the significance test, the process stops.

For each subset of variables we first provided a descriptive measure of their association, and then, via the model selection procedure, we provided the dependence graph and the estimated partial correlation matrix of the selected model. Since all the variables were continuous, the estimated models are Gaussian graphical models (See Højsgaard et al. (2012) for the implementation in the statistical software R). To save space we do not report the estimated partial correlations. Note however that that the results were in all cases very close to the empirical partial correlations.

Section 1: Within-domain BCoS data

As highlighted in the Introduction, the BCoS was designed to assess cognitive performance within 5 different domains. We first examined the relations between the tests within each domain, to determine the within-domain structure when considered in isolation.

Attention and executive function tests. The empirical partial correlation matrix of the attention and executive function tests variables shown in Table 1 reflects the correlation of each pair of variables after taking into account all the remaining ones in the domain. What is noteworthy is that the partial correlations were relatively sparse. Overall performance on the

Apple cancellation task (a non-lateralised measure of spatial selection; Bickerton et al., 2011, in press), correlated with a measure of lateralised asymmetry on the same task (the Apple asymmetry score), measures of extinction and the rule finding and switching task from BCoS. However the Apple asymmetry score, a measure of egocentric neglect (Bickerton et al., 2011), had minimal correlation with the other tests, including the measures of spatial extinction. The scores for left visual and tactile extinction partially correlated, and there was also a partial correlation between the auditory attention test and the rule finding and switching task.

APC	ASY	LVE	RVE	LTE	RTE	AUD	RUL
100.00	**−25.96**	**−14.62**	**−16.76**	**−24.63**	2.14	5.72	**29.86**
	100.00	1.50	−7.40	1.85	−9.52	−7.16	2.51
		100.00	−4.61	**37.10**	−0.09	−0.87	−7.51
			100.00	−10.01	7.83	−5.29	7.86
				100.00	5.70	−1.17	−0.93
					100.00	−4.58	−0.45
						100.00	**25.97**
							100.00

Table 1: Empirical partial correlation matrix of the Attention variables. APC = Apple cancellation (total score); ASY = Apple cancellation page asymmetry; LVE = left visual extinction score; RVE = right visual extinction score; LTE = left tactile extinction score; RTE = right tactile extinction score; AUD = total score on the auditory attention test; RUL = total score on the rule finding and set shifting test (measuring executive functions). In this and all other tables statistically reliable partial correlations are shown in **bold.**

The estimated dependence graph (Figure 1), shows that right extinction scores (for both visual and tactile tests) were isolated from the other variables. Note that these deficits are associated with left hemisphere lesions while the other deficits have greater right hemisphere involvement (see Bickerton et al., 2011, in press). Within the other variables in the attention/executive function domain there was an association between the Apple cancellation task (overall performance) and (a) left visual and left tactile extinction (LVE and LTE), (b) the Apple asymmetry score (ASY) and (c) the rule finding and shifting task. These associations suggest that the overall Apple cancellation score is related to 3 factors: (i) a left spatial asymmetry that is detected under extinction conditions (LVE and LTE); (ii) measures reflecting executive function (RUL); and (iii) a measure of spatial neglect (ASY). Interestingly, once the overall Apple score was taken into account, there was no direct relationship between the left extinction measures (LVE and LTE) and the neglect measure (ASY), suggesting some distinction between extinction and neglect and that extinction does not merely represent ‘mild neglect’ (e.g., Chechlacz et al., 2013; Karnath et al., 2003). Indeed the independent link between the non-lateralised cancellation score (Apple overall cancellation, APC) and extinction suggests that extinction may reflect the ability to select competing targets over and above effects based on the spatial positions of the stimuli. There was also no relation between the spatial bias measures (e.g., LTE and LVE) and performance on the executive rule finding test (RUL), once the overall APC score was taken into account. The deviance of the model was 24.70 with 22 degrees of freedom (p-value 0.31) providing a good fit.

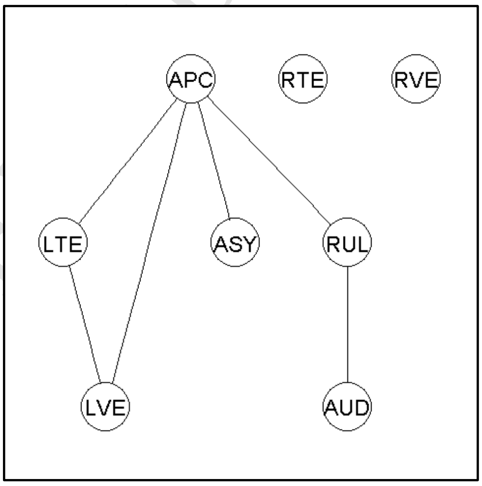


Figure 1: Dependence graph for the Attentional/Executive Function domain. We illustrate how to interpret the graph with one example. In Figure 1 the graph shows that vertex LTE and LVE are separated from vertex ASY by APC, this means that LTE and LVE are conditionally independent of ASY, given APC.

Language tests. The empirical partial correlation matrix for the language variables (Table 2) indicates several features. The Instruction comprehension score (ISC) had generally low partial correlations with the remaining variables which might reflect relatively low sensitivity for this measure and/or that this provides the only ‘pure’ test of language comprehension. Picture naming (PIC) on the other hand was associated with several other variables requiring language production including sentence construction (SCS), sentence reading (SRD) and writing words/nonwords (WWN). The tests requiring spoken production (sentence construction and reading) however were not strongly correlated with written production (WWN), consistent with a dissociation between spoken and written production. The nonword reading test (RNW) was correlated with sentence reading (SRD) and writing (WWN), consistent with nonword reading requiring both speech output and non-lexical phonological processing.

PIC	SCS	SRD	WWN	ISC	RNW
100.00	**36.69**	**32.11**	**23.05**	−0.12	4.31
	100.00	**29.06**	−5.88	10.63	15.08
		100.00	2.31	1.38	**34.35**
			100.00	6.35	**34.74**
				100.00	−4.95
					100.00

Table 2: Empirical partial correlation matrix within the Language domain. PIC = picture naming; SCS = sentence construction score; SRD = sentence reading; WWN = writing words and nonwords; ISC = instruction comprehension; RNW = reading nonwords.

The dependence graph (Figure 2) showed a close association between sentence construction (SCS) and: sentence reading (SRD), picture naming (PIC), nonword reading (RNW) and instruction comprehension (ISC). However once the sentence construction score was known, the measure of comprehension (ISC) was conditionally independent of all the remaining variables. The deviance was 3.04 with 7 degrees of freedom (p-value 0.88), indicating a good fit for the model.

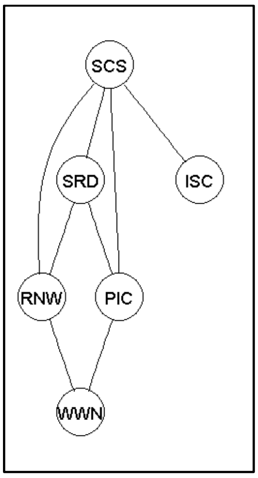


Figure 2: Dependence graph for the analysis within the Language domain.

Memory tests. The empirical partial correlation matrix for the memory test variables is shown in Table 3. There were reliable correlations between (i) two aspects of personal memory orienting - recall of occupation, age, qualifications (PER) and correct report of where the patient is, time and date (TSFR), (ii) the immediate and delayed recall scores (SImF and SDeF) and (iii) the delayed memory tests (recognition, TAR, and SDeF involving recall) and also the test of being oriented in time and space (TSFR).

PER	TSFR	NOS	SImF	TAR	SDeF
100.00	**29.35**	**22.19**	6.45	2.25	−1.17
	100.00	1.74	6.80	**20.22**	15.68
		100.00	0.14	1.31	9.18
			100.00	0.56	**60.52**
				100.00	**31.17**
					100.00

Table 3: Empirical partial correlation matrix of variables within the Memory domain. PER = Personal information recall; TSFR = time and space free recall; NOS = nosognosia; SImF = story immediate free recall; TAR = task recognition; SDeF = story delayed free recall.

The dependence graph for the memory tests is depicted in Figure 3. The most challenging measure of long-term memory, delayed recall (SDeF), was linked to task recognition (TAR), memory for location in space and time (TSFR) and knowledge of why the patient was there

(NOS). Recall of personal information (PER) was linked to memory about the current situation (TSFR), to knowledge of symptoms (NOS) and to immediate recall (SImF). In each case performance depends on good maintenance of information about the current situation (the patient's own situation and also recently presented words). The deviance was 1.98 with 7 degrees of freedom (p-value 0.96), giving no evidence to reject the model. Long-term memory for personal information (PER) was not directly related to long-term delayed recall (SDeF).

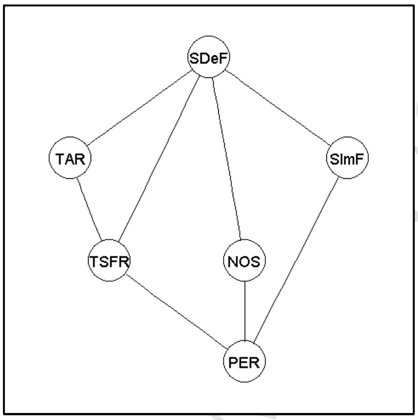


Figure 3: Dependence graph within the Memory domain.

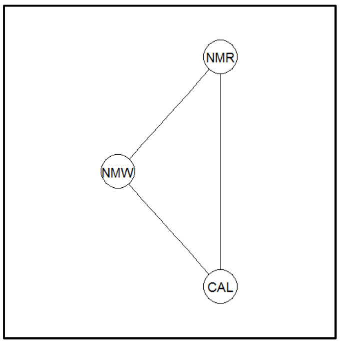
Number processing tests. The empirical partial correlation for the number processing tests (Table 4) indicated some association between all the variables. This was confirmed by the estimated dependence graph (Figure 4) where NMR (number/price reading), NMW (number/price naming) and CAL (calculation) formed a complete graph. This analysis indicates that the number processing tests were highly inter-related when analyzed in a single domain.

NMR	NMW	CAL
100.00	**50.67**	12.46
	100.00	**38.83**
		100.00

Table 4: Empirical partial correlation matrix for the graph analysis within the Number domain. NMR = number reading; NMW = number writing and CAL = calculation performance.

The dependence graph here represents a saturated model with no conditional independences between the variables and therefore the estimated partial correlation matrix is the same as the empirical correlation matrix and the deviance of the model was 0.

Figure 4: Dependence graph for the Number domain.



Praxis tests. The empirical partial correlation matrix for the praxis tests (Table 5) indicated generally reliable partial correlations across the tests with the strongest correlations being between the multi-step object test (MOU) and the complex figure copy (CFC), the complex figure copy and the gesture imitation test (GEI), and between gesture production (GEP) and gesture recognition (GER) and gesture imitation (GEI). The multi-step object test and the complex figure task both involve sequential behaviour. The complex figure and gesture imitation both demand memory for action. The gesture production, recognition and imitation tasks all involve the coding of hand actions.

MOU	GEP	GER	GEI	CFC
100.00	17.24	19.11	7.37	**33.26**
	100.00	**21.09**	**32.95**	−7.25
		100.00	13.53	−2.02
			100.00	**36.12**
				100.00

Table 5: Empirical partial correlation matrix for the Praxis domain. MOU = multi-step object use; GEP = gesture production; GER = gesture recognition; GEI = gesture imitation; CFC = complex figure copy.

Figure 5 shows the dependence graph for the praxis tests. The analysis indicated close inter-relations between the three gesture tasks (GER, GEP and GEI), and between the tasks dependent on stored gesture knowledge (GER and GEP) and the multi-step object test (MOU). The complex figure copy (CFC) was linked to the multi-step object use test and the gesture imitation test, perhaps reflecting its dependence on both multi-step planning and visual memory (see Bonivento et al., 2013, for evidence of the relations between visual memory and the ability to imitate meaningless gestures). The deviance of the fitted model was 4.45 with 3 degrees of freedom (p-value 0.22), and there was no evidence to reject the model.

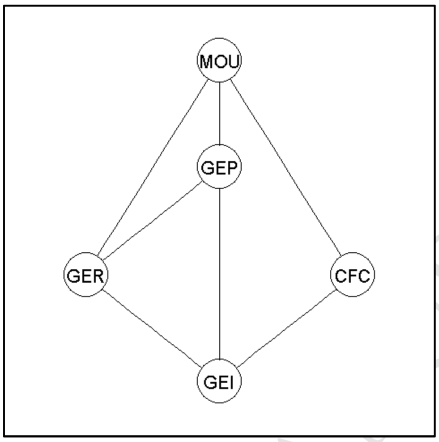


Figure 5: Dependence graph for Praxis tests.

Discussion

The results of the within-domain analyses generally show patterns of connectedness between tests designed to tap different parts of the cognitive system. For example, in the attention domain there is a separation between tests where a left spatial bias is evidence (e.g., the Apples Asymmetry and the left extinction tests) and measures of executive function, and both are distinct from spatial attention biases associated with left hemisphere damage (right-side extinction). In the language domain the measure of language comprehension (Instruction comprehension) separated from tests requiring phonological output processes, and tests requiring phonological production differed from those involving written production. In the memory domain no direct edges connected the immediate free recall measures and the delayed recognition measures, once delayed recall was taken into account, consistent with the involvement of distinct immediate and longer-term memory processes which might draw on common retrieval processes (tapped by the delayed recall measure). In the number domain reading, writing and calculation were highly inter-linked, suggesting a dependence on a common representation for number (though see below for an alternative proposal following the across-domain analyses). In the praxis domain the gesture production, recognition and imitation tasks were closely linked while there were linkages between the multi-step object and complex figure tasks - consistent with their both being dependent on action sequencing. These results are broadly consistent with cognitive neuropsychological models in each domain (e.g., Ellis & Young, 1988). In Section 2 we go on to evaluate if these within-domain relationships are maintained when the full pattern of variance is taking into account, involving performance on tasks in other domains.

Section 2: Across-domain BCoS data

In high dimensional settings, graphical models can be particularly useful because they allow visual inspection of the structural relations between sets of variables. In Section 2 we analysed the relations between all the subtests in the BCoS in an interaction model[3] following the procedure introduced by Edwards et al. (2010) (and implemented in the 'gRapHD' package; Abreu et al., 2010, in the statistical software R). This procedure involves finding an initial structure (the minimal forest) and then performing stepwise model selection starting from that. Stepwise selection starts from the previously found forest using forward search by adding edges that improve the model (using BIC). The selection stops if there is no such edge available. When extra factors are entered into the analysis, is it possible that some of the edges in the original domain-specific models may disappear because partial correlations between the variables are absorbed into correlations with the additional factors.

Figure 6 shows the graph obtained after performing the stepwise selection on all the

[3] A stepwise procedure starting from the independence graph, as performed in the previous section, is computationally impractical.

BCoS variables plus also the measures of initial motor function (Barthel index) and affect (HADS), with each domain labelled by a different colour[4]. The figure illustrates two points. At a general level, many of the nodes making up each domain remained integrated, supporting the reality of the different cognitive domains. On the other hand, there were substantive changes in the details of the models within each domain. Notably, over 50% of the edges between the tests within each domain disappeared in the domain-general analysis. In the domain-general analysis there was separation within the following domains: (i) memory (where the personal memory and anosagnosia tests separated from the other memory tests), (ii) attention and executive functions (where the auditory attention test linked more strongly to aspects of number processing, language and memory, while right extinction (RVE and RTE) remained distinct from left extinction and neglect (LVE, LTE, ASY)) and (iii) praxis, where the gesture recognition test linked to auditory attention more than the other gesture tasks. In addition, the cognitive test scores were separated from the affective measures (HADS) and only connected to the Barthel score in relation to the complex figure copy, which likely carries a motor control component. This result confirms that cognitive problems after stroke can be distinct from problems in affect and are unlikely to reflect a general deficit reflecting the severity of the stroke.

Within the new analysis there were several interesting, new across-domain links:

1) Number/price reading (NMR) was connected with three language tests - sentence reading (SRD), sentence construction (SCS) and picture naming (PIC), consistent with all the tests depending on spoken word production.
2) Number/price writing (NMW) was connected with the language writing test (WWN), consistent with both requiring the output of written symbols.

[4] We included the Barthel and HADS measures here since it is important to rule out that any changes in cognition did not reflect factors such as depression or the impact of poor motor function (e.g., for the measures of apraxia).

3) The complex figure copy (CFC) was connected with two of the attention and executive function tests - apple cancellation (APC) and rule finding and set switching (RUL) – along with other tests of praxis (MOU, GEI), suggesting that the test involves both praxic and attentional components (e.g., the scanning and switching of spatial attention and the maintenance of a spatial representation; see Chechlacz et al., 2014, for further evidence on the neural basis of complex figure copying).
4) The auditory attention test (AUD) connected with nodes from all other domains. From Figure 6 we can see that the node for this test (AUD) has 11 connecting edges; picture naming (PIC) has 10 and number/price reading (NMR) and complex figure copy (CFC) both have 7 connecting nodes. The links between picture naming and the other tests likely reflect the demands on spoken language in a number of the assessments. However, the cross-domain links found for the auditory attention and complex figure tests suggests that these might be useful markers of impairments across different domains, and might be adopted as initial tests where there is limited time to assess patients.

These data with the auditory attention test are consistent with it having several components – working memory for the target and distractor words, sustained attention across the trial blocks, and response inhibition to prevent erroneous responses to distractors (see Humphreys et al., 2012). These cognitive processes (working memory, sustained attention, response inhibition) will modulate many other tasks. It is noticeable also that the measure of understanding the task instructions (ISC), though putatively a language task was not connected to the other language tests, but did link to the measure of auditory attention and long-term memory (SDeF). The result suggests that comprehension of task instructions may be as reliant on sustained attention and the ability to consolidate information in long-term memory as upon language abilities per se (see Francis et al., 2003).

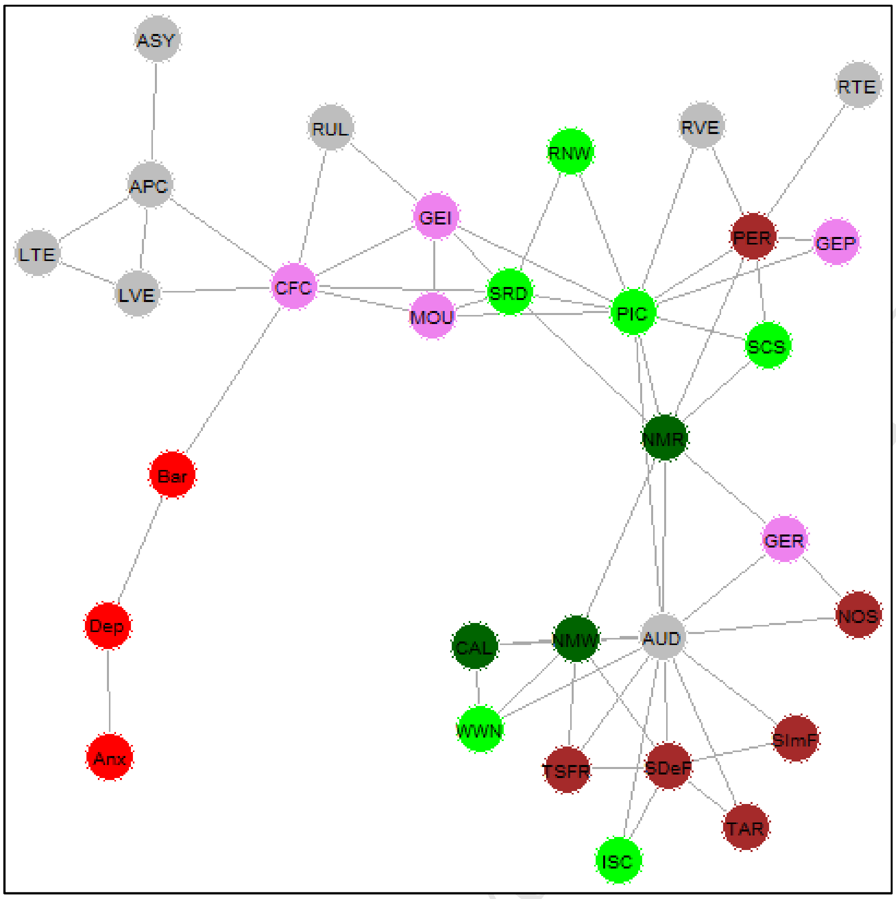


Figure 6: Stepwise selection on all variables (red: Clinical behavioural information variables; grey: Attention and Executive attention test variables; green: Language test variables; dark green: Number skills variables; brown: Memory test variables; violet: Praxis variables.

Discussion

Graphical models allow us to understand complex associations between different variables. We have shown here that graphical model analyses can reveal the structure of the tasks included in the BCoS by demonstrating conditional independence relations between different groupings of the variables. One set of relations was evident in the within-domain analyses. Interestingly, however, there were substantial changes in the underlying relations between the tests when all the measures were considered together to take domain-general factors into account.

Domain-specific analyses

Our first analysis revealed the structure between test scores within each of the sub-domains of the BCoS, when considered in isolation. The results indicated that the overlap in test scores varied across the domains. For example, the number processing tasks (reading, writing and calculation) were all interlinked suggesting substantial inter-relations even though the reading and writing tasks required number transcription without necessarily requiring access to semantic information about number magnitude (unlike calculation). The subsequent full test analysis is helpful here since this showed that performance on the number tasks was closely related to that on the Auditory Attention task. This in turn suggests that holding information in working memory may be critical to performance on the number tasks (even number reading, where several digits may have to be maintained whilst being recorded into a phonological form).

In the praxis domain, the complex figure task related to the gesture imitation task and the multi-step task, all of which require sequences of action to be maintained and produced. The gesture recognition and production tasks were unrelated to this component however, and deficits in these sub-tests more likely reflect access to representations for the recognition and production of single actions.

In the memory domain close inter-connectivity was found between the verbal recall tasks (delayed and immediate) and the recall of personal information (PER) and information about why the patient was in the hospital (NOS). This fits with all of these tasks demanding verbal retrieval processes. Delayed recall was separately related to task recognition and recognition in time and space. These last tasks appear to depend more on memory consolidation processes, required to recall after a delay, to recognise items used in the tests and to which day or month it was (a task that could not be cued by recognition of the environment, unlike recognising being in the hospital).

There were also sub-divisions within the language domain. One cluster of inter-linked tests was concerned with reading (words and nonwords), writing nonwords and picture naming. On the other hand, reading (words and nonwords again), picture naming and sentence construction were also linked. These distinct clusters raise the question of what were common underlying factors. Here the across-domain analyses are again useful because, in these cases, the instruction memory linked to the auditory attention task and there was a separate link to auditory attention for the task requiring the writing of words and nonwords. For the auditory attention task incoming target words must be consolidated into verbal working memory, and this short-term consolidation process may be important for understanding instructions from verbal sentences. The on-line maintenance of phonological representations for the target words, however, may be more critical for writing words and nonwords. Thus, we distinguish between on-line holding of phonological representations (writing words and nonwords) and the consolidation of input phonological representations (comprehending verbal instructions). On the other hand, the lack of relation between instruction comprehension and nonword writing in the overall model, suggests that some form of common output phonology was not a critical factor. It may be that writing depends on small phonological units (especially for nonwords) while instruction comprehension relies on larger units, and this leads to the lack of relation between instruction comprehension and writing in the overall model.

The distinction between instruction comprehension and picture naming, reading and sentence construction in the overall model can more clearly be linked to the distinction between input and output phonological representations (Howard & Nickels, 2005). Instruction comprehension, linked to auditory attention and the consolidation of auditory words, may be strongly weighted for phonological input representations while picture naming, reading and sentence construction tasks weight output phonology.

The relations between the tasks within each sub-domain can broadly be conceptualised in terms of cognitive neuropsychological models of cognition (Ellis & Young, 1988) – for example discriminating between immediate recall and recognition, the sequencing and recognition of action, input and output phonology.

Domain-general analyses

In contrast to the domain-specific results, the domain-general analyses revealed that several tests can account for data outside their specific domain. Notably, the complex figure copy test along with the auditory attention task has multiple connections outside of their originally designated domains. The complex figure task links not only to measures of praxis but also to measures of attention (particularly spatial attention indices on the Apples cancellation task). This fits with a recent lesion-symptom mapping study of Chechlacz et al. (2014). These authors took conducted a voxel-based morphological analysis of the relations between brain lesions and performance on the complex figure task. They found distinct lesion sites were associated with contrasting measures – whether errors were lateralised on one side of the figure (linked to posterior parietal damage), whether there was poor positioning of local features across the entire figure (linked to more ventral visual lesions) and so forth. The results suggest that several factors contribute to performance on this test and, consistent with this, we show that scores on the complex figure are related to attention as well as praxis in our overall analysis.

Consider also the auditory attention test. While categorized into the attention and executive function domain, performance on this test is also related to patients' memory abilities within our overall model (see also the gesture recognition test). This is not surprising given that the auditory attention test was designed to measure several factors including working memory and sustained attention as well as the ability to select targets and reject distractors.

The overall analysis, then, helps to show the underlying factors that contribute to the performance of such multi-faceted tests. Note that this highlights that graphical modelling reveals directly the relations between tests, while our conclusions about the relations between underlying cognitive components is deduced (indirectly) from the structural relations between the tests. Thus we should not conclude that the concept of auditory attention is necessarily closely related to memory, but rather that this is the case in the BCoS battery (where the test was designed to assess working memory and not just the ability to select targets and reject distractors). We presume that different tests will ‘weight’ contrasting conceptual components to varying degrees. The analysis does reveal one other important thing though. This is that the sub-tests which link across several domains might serve well as initial probes of performance, if a clinician wishes to gain a ‘quick and dirty’ analysis of cognition before setting off to track-down which more specific processes might (also) be impaired – through subsequent analyses at the sub-domain level. In the auditory attention test here, the focusing on more specific sub-domains can be guided by the separate measures of selection, working memory and sustained attention. The complex figure test may provide some initial indication of poor spatial attention (neglect), alongside the problem in drawing construction.

A further critical aspect of the overall model analysis was that the connectivity evident within the domains was greatly reduced when variation linked to the tests in other domains was taken into account. For example, once co-variation in working memory consolidation linked to the auditory attention task was extracted, then the instruction comprehension measure became decoupled from other language tasks – in these other language tasks there may have been some component of working memory but it was less strongly weighted (e.g., in sentence construction). The results suggest that the domain-specific models linked the language tests through co-dependence on working memory. However, the relations between these tests and the instruction comprehension measure reduced when a test accounting for more of the variance associated with working memory (the auditory attention test) was introduced.

The full model analysis reveals that associations can be strengthened or weakened or even reversed – a phenomenon referred to in statistics as Simpson's paradox (Simpson's 1951) – when the variance linked to other factors is taken into account. There are clinical implications. Notably we suggest that clinical assessment should incorporate domain-general as well as domain-specific assessments, precisely to tease apart the relative interplay between domain-specific and domain-general processes in a given patient. For example, our results indicate that a clinician should be cautious in making an association between poor language (e..g, on picture naming) and poor comprehension of instructions, concluding that the language impairment is responsible for the poor comprehension. Our cross-domain analysis indicates that instruction comprehension could link to impaired working memory rather than poor language per se. Indeed, given the presence of co-varying domain-general and domain-specific deficits, then the domain general deficits might be the principle target for rehabilitation given that any improvement in domain general processing may generalise more. Consistent with this, Francis et al. (2003) provided evidence that training working memory improved sentence comprehension in patients. Brownsett et al. (2014) also reported that the activation of frontal brain networks concerned with executive attentional control was associated with the communicative abilities of aphasic patients – over and above effects of lesion size. Brownsett et al. propose that damage to the frontal executive network is predictive of the degree of language impairments suffered by aphasic patients, and such domain general problems need to be taken account of alongside the domain-specific impairments in language. The present results concur with this and indicate that cognitive processes assessed using the auditory attention test of BCoS likely contribute to any language impairments in the patients – we suggest that these processes include working memory consolidation (see above) but perhaps also sustained attention and the ability to suppress irrelevant distractors. We conclude that it is important to document the cognitive profile of individual patients, as is done by the BCoS battery, so that their domain-general impairments are noted along with any domain-specific problems.

Within cognitive neuropsychological models of cognition, the stress has been on the input-output transformations that can operate between processing modules in a given domain – for example how input lexical information is mapped onto stored semantic knowledge. The role of domain-general processes has been less easy to conceptualise (e.g., the role of working memory and sustained attention), and perhaps for this reason many of the standard cognitive neuropsychological tests (e.g., the PALPA, Kay, Coltheart & Lesser, 1992) do not assess such processes alongside the domain-specific transformations. We believe this can be misleading, given that domain-specific problems co-vary with the presence of the domain-general impairment (e.g., Bickerton et al., in press; Brownsett et al., 2014; Corbetta et al., 2015) even with lesion size taken account of. The cognitive profiling approach of the BCoS provides one solution.

The relations between what we are terming domain-general processes and cognitive resources is also worth considering. Shallice (1988), amongst others, has noted the importance of taking account of resource deficits in neuropsychological patients along with any domain-specific processes. Here the idea of ‘resource’ can be operationized in terms of the average proportion of neurons functioning normally in a particular sub-system, which are needed to produce a given level of performance (Shallice, 1979), with the effects of a brain lesion being to reduce the proportion of operational neurons. It may be indexed by abnormal effects of task difficulty in a given patient. Our argument, for the importance of domain-general as well as domain-specific components, differs from this however. Notably we highlight the role of cognitive processes that span multiple domains (e.g., sustained attention, working memory) but do not simply reflect task difficulty or the proportional loss of neurons in critical domain-specific regions. Corbetta et al. (2015) propose that the domain-general deficits may stem from white matter damage to key ‘cross road’ regions containing multiple fibre tracts which support cross-talk across cognitive domains. Here critical lesions may not necessarily be large but would affect the critical pathways supporting this cross-talk. Indeed in their data analysis Corbetta et al. extracted out the effects of lesion size.

Bickerton et al. (in press) also report that functional recovery in patients relates to the presence of co-occurring cognitive deficits measured using the BCoS battery, so that (e.g.) recovery is worse if patients had poor executive attention alongside a memory deficit (see also Brownsett et al., 2014). The effect of the co-occurring problem again arose when effects due to lesion size were extracted. These results suggest that simple loss of ‘resource’, in terms of the overall proportion of brain tissue affected, is less critical here than the loss of additional support structures (domain-general operations) that underpin domain-specific cognitive operations.

Other forms of multivariate analysis, for example to extract underlying principal components, have been used in the analysis of brain lesion data (correlating the weighting on the given component for an individual against their lesion results; Corbetta et al., 2015; Chechlacz, Rotshtein & Humphreys, 2014; Verdon et al., 2010). Here it has been argued that the critical conceptual component can be localised in the area(s) where a correlation with the lesion is shown. One question for future work is whether the graph analysis being proposed here can also be used to guide our understanding of the neural basis of cognition, for example by being incorporated into lesion-symptom mapping studies. As we have noted, the graph analyses we present are focused on the relations between tests rather than on the underlying conceptual components, and to some degree it may be of limited help to localise a test using lesion-symptom mapping procedures. However it might be of interest to attempt to localise some of the edges in a model especially where the edges can more clearly be linked to conceptual processing components. To do this we would need to derive information about how well an individual correlates with the group as an index of the strength of the 'edge' for an individual, which can then be used as a regressor in modelling lesion-symptom relations. In this way we may be able to extract the neural correlates of a particular 'edge'. This proposal awaits future research.

One final point is that our overall analysis indicated that the cognitive measures were largely independent of variation in anxiety, depression and initial motor function (Barthel index), consistent with the cognitive problems experienced by patients not being determined by low affect (see also Nys et al., 2005). The data suggest that cognitive problems following stroke can be dissociated from poor affect and initial motor function, but that it is helpful to take the presence of domain-general cognitive problems into account.

Acknowledgements

This work was supported by grants from the Stroke Association and the National Institute of Health Research (Oxford cognitive health CRF).

References

Abreu, G. C. G., Labouriau R. and Edwards, D. (2010). High-Dimensional Graphical Model Search with the gRapHD R Package. Journal of Statistical Software, **37**(01).

Bickerton, W-L., Demeyere, N., Francis, D., Kumar, V., Remoundou, M., Bahrami Balani, A., Harris, L., Williamson, J., Lau, J.K., Samson, D., Riddoch, M.J. and Humphreys, G.W. (in press) The BCoS cognitive profile screen: Utility and predictive value for stroke.

Neuropsychology.

Bickerton, W-L., Riddoch, M.J., Samson, D., Bahrami Balani, A., Mistry, B. and Humphreys, G.W. (2012) Systematic assessment of aprxia and function predictions from the Birmingham Cognitive Screen. Journal of Neurology, Neurosurgery and Psychiatry, **83**, 513-521.

Bickerton, W-L., Samson, D., Williamson, J. and Humprheys, G.W. (2011) Separating forms of neglect using the Apples Test: Validation and functional prediction in chronic and acute stroke. Neuropsychology, **25**, 567-580.

Bonivento, C., Rumiati, R.I., Biasutti, E. & Humphreys, G.W. (2013). The role of the Basal Ganglia in Action Imitation: Neuropsychological evidence from Parkinson's disease Patients. Experimental Brain Research, **224**(2), 211-20.

Brownsett, S.L.E., Warren, J.E., Geranmayeh, F., Woodhgead, Z., Leech, R. & Wise, R.J.S. (2014) Cognitive control and its impact on recovery from aphasic stroke. Brain, **137**, 242-254.

Chechlacz, M., Novick, A., Rotshtein, P., Bickerton, W-L., Humphreys, G.W. & Demeyere, N. (2014) The neural substrates of drawing: A VBM analysis of constructional, hierarchical and spatial representation deficits. Journal of Cognitive Neuroscience, **26**, 2701-2715.

Chechlacz, M., Rotshtein, P. & Humphreys, G.W. (2014) Neural substrates of Corsi Block span: lesion-symptom mapping analyses in relation to attentional competition and spatial bias. Neuropsychologia, **64**, 240-251.

Chechlacz, M., Rotshtein, P., Hansen, P. C., Deb, S., Riddoch, M. J., & Humphreys, G. W. (2013). The central role of the temporo-parietal junction and the superior longitudinal fasciculus in supporting multi-item competition: Evidence from lesion-symptom mapping of extinction. Cortex, **49**(2), 487-506.

Corbetta, M. et al. (2015) Common behavioural clusters and subcortical anatomy in stroke. Neuron, **85**, 927-941.

De Haan, R. J., Limburg, M., Van Der Meulen, J. H., Jacobs, H. M. and Aaronson, N. K. (1995) Quality of life after stroke. Impact of stroke type and lesion location. Stroke, a journal of cerebral circulation, **26**(3), 402-408.

Demeyere, N., Slavkova, E., Riddoch, M.J. & Humphreys, G.W. (in press) The Oxford Cognitive Screen (OCS): Validation of a stroke-specific short cognitive screening tool. Psychological Assessment.

Edwards, D. (2000) Introduction to Graphical Modelling. New York: Springer.

Edwards, D., De Abreu, G. C. and Labouriau, R. (2010). Selecting high-dimensional mixed graphical models using minimal AIC or BIC forests. BMC Bioinformatics, **11**, 18.

Ellis, A.W. & Young, A.W. (1988) Human cognitive neuropsychology. London: Psychology Press.

Fillingham, J.K., Sage, K. & Lambon-Ralph, M. (2005) Treatment of anomia using errorless versus errorful learning: Are frontal executive skills and feedback important? Int. J. Lang. & Commun. Disorders, **40**, 505-523.

Gibbs, J. W. (1902). Elementary principles in statistical mechanics. Woodbridge, CT: Ox Bow Press.

Højsgaard, S., Edwards, D. and Lauritzen, S. L. (2012). Graphical Models with R. New York: Springer.

Howard, D. & Nickels, L. (2005) Separating input and output phonology: Semantic, phonological, and orthographic effects in short-term memory impairment. Cognitive Neuropsychology, **22**, 42-77.

Humphreys, G., Bickerton, W.L., Samson, D. and Riddoch, M.J. (2012) BCoS Cognitive Screen. Hove: Psychology Press.

Jørgensen, H. S., Nakayama, H., Raaschou, H. O. and Olsen, T. S. (1995). Intracerebral hemorrhage versus infarction: stroke severity, risk factors, and prognosis. Annals of neurology, **38**(1), 45-50. 49

Jørgensen, H. S., Nakayama, H., Reith, J., Raaschou, H. O. and Olsen, T. S. (1997). Stroke recurrence: Predictors, severity, and prognosis. The Copenhagen Stroke Study. Neurology. **48**(4), 891-895.

Karnath H.O., Himmelbach M., and Kuker W (2003.) The cortical substrate of visual extinction. Neuroreport, **14** (3):437-442.

Kay, J., Coltheart, M. & Lesser, R. (1992) PALPA: Psycholinguistic Assessments of Language Processing in Aphasia. London: Psychology Press.

Lauritzen, S. L. (1996). Graphical models. Oxford: Oxford University Press.

Mahoney F. and Barthel D. (1965). Functional evaluation: the Barthel Index. Md. Med. J. **14**, 61–65.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A. & Wager, T.D. (2000) The unity and diversity of executive functions and their contributions to complex ‘frontal lobe’ tasks: A latent variable analysis. Cognitive Psychology,**41**, 49-100.

Nys, G.M.S., van Zandvoort, M.J.E., de Kort, P.L.M., van der Worp, H.B., Jansen, B.P.W., de Haan, E.H.F. and Kappelle, L.J. (2005) The prognostic value of domain-specific cognitive abilities in acute first-ever stroke. Neurology, **64**, 821-827.

Schwarz, G. (1978). Estimating the dimension of a model. Ann Math Stat. **6**, 461-464.

Simpson, E.H. (1951) The interpretation of interaction in contingency tables. Journal of the Royal Statistical Society, **B13**, 238-241.

Shallice, T. (1979) Case-study approach in neuropsychological research. Journal of Clinical Neuropsychology **1**, 183-211.

Shallice, T. (1988) From neuropsychology to mental structure. Cambridge: Cambridge University Press.

Sims, N.R. and Muyderman, H. (2009) Mitochondria, oxidative metabolism and cell death in stroke. Biochimica et Biophysica Acta. **1802** (1), 80–91.

Verdon, V., Schwartz, S., Lovblad, K-O., Hauert, C-A. & Vuilleumier, P. (2010) Neuroanatomy of neglect and its functional component: A study using voxel-based lesion-symptom mapping. Brain, **133**, 880-894.

Whisnant J.P. (1996). Effectiveness versus efficacy of treatment of hypertension for stroke prevention. Neurology **46** (2), 301–7.

Whittaker, J. (1990). Graphical models in applied multivariate statistics. Chichester: Wiley.

Wold, H. (1954). Random normal deviates. Cambridge: Cambridge University Press. 50

Wolf P.A., Abbott R.D. and Kannel W.B. (1987). Atrial fibrillation: a major contributor to stroke in the elderly. The Framingham Study. Arch. Intern. Med. **147** (9), 1561–4.

Wright, S. (1921). Correlation and causation. Journal of Agricultural Research. **20**, 557-585.

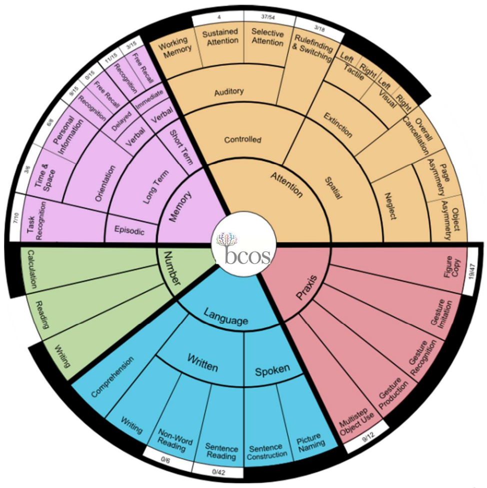
Wyller, T, B., Sødring, K, M., Sveen, U., Ljunggren, A, E. and Bautz-holter, E. (1997), Are there gender differences in functional outcome after stroke? Clinical Rehabilitation **11** (2), 171-9.

Zigmond, A. S. and Snaith, R. P. (1983). The hospital anxiety and depression scale. Acta Psychiatrica Scandinavica. **67** (6), 361–370.

APPENDIX 1

BCoS is design to provide a cognitive profile on a patient covering 5 areas of cognition and including both domain-specific and domain-general processes. The administration takes around 1 hour.

Areas	**Tests**	**Domain classification**
Attention and executive function	Auditory attention test	**General** (working memory, sustained attention, target selection)
	Rule finding and concept shifting	**General** (rule finding, set shifting)
	Apple cancellation	**Specific** (spatial orienting)
	Visual extinction	**Specific** (visual selection)
	Tactile extinction	**Specific** (tactile selection)
Language	Picture naming	**Specific** (object recognition, naming)
	Sentence construction	**Specific** (syntax)
	Instruction comprehension	**Specific** (sentence comprehension)
	Sentence reading	**Specific** (word recognition, naming)
	Reading nonwords	**Specific** (nonlexical processing)
	Writing words and nonwords	**Specific** (lexical and nonlexical production)
Memory	Orientation	**Specific** (contextual orienting)
	Story recall and recognition	**Specific** (immediate and long-term verbal memory)
	Task recognition	**Specific** (long-term visual memory)
Number skills	Number reading	**Specific** (number recognition, naming)
	Number writing	**Specific** (number production)
	Calculation	**Specific** (arithmetic operations)
Praxis	Complex figure copy	**General** (working memory, spatial representation)
	Multi-step object use	**General** (recognition, sequencing, planning)
	Gesture production	**Specific** (gesture retrieval)
	Gesture recognition	**Specific** (gesture production)
	Imitation	**Specific** (nonlexical transcription)



The cognitive profile for BCoS, used for case reporting. Each domain is assigned a separate colour in the pie chart. A black edge indicates that a patient has performed at a normal level. A white edge indicates a score falling below the level for age- and education-matched controls. Where the edge segment is missing, the patient has not been tested. Here the patient has deficits in memory but some preserved aspects of language and praxis.

APPENDIX 2

BCoS Task descriptions

The BCoS is designed to provide an overall ‘cognitive profile’ for patients after brain injury, covering 5 areas of cognition: (i) attention and executive function; (ii) language; (iii) memory; (iv) number processing and (v) praxis/skilled action. The sub-tests aim to measure both domain-specific abilities (primarily affecting just one of the areas listed above) and domain-general processes (processes that impact on abilities outside the targeted area – an example being executive functions which can affect language, memory etc.). The tests are designed in order to gain maximal inclusion for patients whilst also being time efficient in their delivery. Time efficient delivery is established by having sub-tests generate separate measures linked to distinct cognitive functions (e.g., see the Auditory attention task). Inclusivity is gained by making the tests ‘aphasia-‘ and ‘neglect-friendly’. Thus for non-language tests the BCoS uses high frequency words and forced-choice testing procedures where possible – so that aphasic patients can still generate responses. For tests not aimed at assessing spatial attention, the stimuli are centred on the page and multi-modal presentations are used.

1. ATTENTION AND EXECUTIVE FUNCTIONS

1.1. Auditory attention task. The task consists of 6 high frequency words presented 9 times each. Half are target words to respond to, half are distractors words which have to be ignored. Each target word (“no”, “hello”, “please”) has a closely related distractor (“yes”, “goodbye”, “thanks”). The words are presented in random order, each being preceded an equal number of times by a 2 sec, 3 sec or 4 sec. silence gap. The task is performed in 3 blocks, providing a measure of how well patients can *sustain their atte*ntion across the blocks. It also measures whether the patients can selectively attend to the target words and prevent themselves from responding to the related distractors (*target selection).* In addition, patients are asked to recall the target and distractor words at the end of the task, providing a measure of whether they can store items in memory over the short-term when they are engaged in another activity (a measure of *working memory*).

1.2. Rule finding and concept switching. Each stimulus consists of a grid made of 6 columns and 6 lines with 32 grey cells, 2 red and 2 green. The task is to learn to predict the movement of a black marker across the grid. The marker moves in a lawful manner but then switches the rule by which it is operating either along a single dimension (position) or across dimensions (position to colour). The task *measures the ability to find an abstract rule and to switch the rule across stimuli within and across dimensions.*

1.3. Apple cancellation. The task consists of an A4 sheet presented in landscape orientation containing complete apples along with distractors which are apples with a left or right part missing. *Egocentric* neglect is measured by whether patients miss targets (whole apples) on one side of the page. *Allocentric* neglect is measured by whether patients make false positive responses by cancelling a 'bitten apple' distractor where the bite is taken from one side.

1.4. Visual extinction. The task consists of 4 unilateral left visual stimuli (finger wiggles by the examiner), 4 unilateral right and 8 bilateral items. Performance is scored according to whether unilateral stimuli are missed (a measure of neglect or a field defect), and whether there is a spatially selective drop in detection on one side when two relative to one stimulus is presented (a measure of extinction).

1.5. Tactile extinction. The task consists of 4 unilateral left, 4 unilateral right and 8 bilateral items. Performance is scored as with the test of visual extinction.

2 LANGUAGE

2.1. Picture naming. The task uses 14 grey shaded hand drawings, half living and half non-living. Half of the items' have a long name in English (6 to 9 letters) and half a short name (3 to 5 letters).

2.2. Sentence construction. The participant sees a photograph of a person carrying out an action and is given two written word. The task is to construct a sentence which describes what the person in the photograph is doing using the two written words.

2.3. Instruction comprehension. This is an index based on the clinical judgement of the examiner, who is asked to rate how well the patient understands the instruction on 4 target tasks and on the number of times the instruction has to be repeated.

2.4. Sentence reading. The task consists of two sentences including both regular and exception words, as well as suffixed and prefixed words.

2.5. Reading nonwords. There are 6 pronounceable nonwords, 5or 6 letters long.

2.6. Writing words and nonwords. The items consist of 4 familiar words (2 regular, 2 exception) and one nonword.

3 MEMORY

3.1. Orientation. The task assesses access to personal information (semantic autobiographic knowledge), orientation in time and space and awareness of deficits (nosognosia).

3.2. Story recall and recognition. The story consists of 15 segments that have to be recalled immediately then after a delay. Each recall test is followed by tests of recognition.

4. NUMBER SKILLS

4.1. Number/price/time reading. The items consist of 3 complex numbers (with units of hundreds and thousands, additive and multiplicative relations, and embedded zeros), 3 prices (all in pounds and pence) and 3 times (digital representation of hours and minutes). The price and time questions are aimed to provide functional measures of the processing of numbers in everyday situations.

4.2. Number/price writing. The items are of the same kind as for the number and price reading task.

4.3. Calculation. Four complex calculations are presented, one addition, one subtraction, one multiplication and one division.

5. PRAXIS

5.1. Complex figure copy. The figure to copy contains a middle structure and additional structures to the left and right. The number of elements to the left and right are equated to balance the probability of left and right neglect. The scoring measures organisation of the figure and associated constructional apraxia as well as the presence of visual neglect.

5.2. Multi-step object use. The task requires the patient to perform a sequence of actions with 2 objects (a battery and a torch) in order to carry out an instruction (light the torch).. Scoring discounts problems due to motor problems/hemiplegia.

5.3. Gesture production. With the least affected hand, the patient is requested to produce 6 actions, 3 intransitive (communicative gestures) and 3 transitive (object-oriented) actions, on verbal command.

5.4. Gesture recognition. The patient is requested to recognise 6 actions, 3 intransitive and 3 transitive actions, that are acted out by the examiner. The patient is asked to make a choice from 4 stimuli for their response, and the stimuli are presented as written words and read aloud by the examiner.

5.5. Imitation. Four meaningless gestures are presented. Two involve a sequence of 2 hand positions in relation to the head and 2 involve a single finger position. The patient is asked to mimic with the least affected hand..